BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 7: e112228
doi: 10.3897/biss.7.112228

OPEN ACCESS

Conference Abstract

An Australian Model of Cooperative Data Publishing to OBIS and GBIF

Katherine Tattersall‡, Peggy Newman§, Sachit Rajbhandari‡, Dave Watts‡, Mahmoud Sadeghi‡

‡ CSIRO, Hobart, Australia
§ CSIRO, Melbourne, Australia

Corresponding author: Katherine Tattersall (katherine.tattersall@csiro.au)

Received: 05 Sep 2023 | Published: 07 Sep 2023

Citation: Tattersall K, Newman P, Rajbhandari S, Watts D, Sadeghi M (2023) An Australian Model of Cooperative Data Publishing to OBIS and GBIF. Biodiversity Information Science and Standards 7: e112228. https://doi.org/10.3897/biss.7.112228

Abstract

The Australian Commonwealth Science and Industrial Research Organisation (CSIRO) hosts both the Australian Ocean Biodiversity Information System (OBIS) and Global Biodiversity Information Facility (GBIF) nodes within the National Collections and Marine Infrastructure (NCMI) business unit. OBIS-AU is led by the NCMI Information and Data Centre and publishes marine biodiversity data in the Darwin Core (DwC) standard via an Integrated Publishing Toolkit (IPT), with over 450 marine datasets at present. The Australian GBIF node is hosted by a separate team at the Atlas of Living Australia (ALA), a national-scale biodiversity analytical and knowledge delivery portal. The ALA aggregates and publishes over 800 terrestrial and marine datasets from a wide variety of research institutes, museums and collections, governments and citizen science agencies, including OBIS-AU. Many OBIS-AU published datasets are harvested and republished by ALA and vice-versa.

OBIS-AU identifies, performs Quality Control and formats marine biodiversity and observation data, then publishes directly to the OBIS international data repository and portal, using GBIF IPT technology. The ALA data processing pipeline harvests, aggregates and enhances datasets from many sources with authoritative taxonomic and spatial reference data before passing the data on to GBIF. OBIS-AU and ALA are working together to ensure that the publication pathways for any datasets managed by both (with potential for duplication of records and incomplete metadata harvests) are rationalised and that a

© Tattersall K et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

single collaborative workflow across both units is followed for publication to GBIF. Recently, the data management groups have established an agreement to cooperatively publish marine data and eDNA data. OBIS-AU have commenced publishing datasets directly to GBIF with ALA endorsement.

We present the convergent evolution of OBIS and GBIF data publishing in Australia, adaptive data workflows to maintain data and metadata integrity, challenges encountered, how domain expertise ensures data quality and the benefits of sharing data skills and code, especially in publishing eDNA data types in DwC (using the DNA-derived data extension) and exploring the new CamTrap Data Package using Frictionless data. We also present the work that both data groups are doing toward adopting the GBIF new Unified Data model for publishing data. This Australian case study demonstrates the strengths of collaborative data publishing and offers a model that minimises replication of data in global aggregators through the development of regional integrated data publishing pipelines.

Keywords

biodiversity, observations, marine, terrestrial, Australia, Atlas of Living Australia, OBIS-AU, DwC, Camera Trap, eDNA

Presenting author

Katherine Tattersall

Presented at

TDWG 2023

Acknowledgements

We acknowledge the use of the Atlas of Living Australia, (https://ror.org/018n2ja79). The Atlas of Living Australia is made possible by contributions from its partners. It is supported by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS) and is hosted by CSIRO. We also acknowledge the use of the CSIRO Marine National Facility (https://ror.org/01mae9353) and the Ocean Biodiversity Information System, Intergovernmental Oceanographic Commission of UNESCO, www.obis.org, in undertaking this work. We acknowledge GBIF: The Global Biodiversity Information Facility (2023) *What is GBIF?* Available from https://www.gbif.org/what-is-gbif 30 June 2023.

Hosting institution

Commonwealth Science and Industrial Research Organisation

Conflicts of interest

The authors have declared that no competing interests exist.